p-12

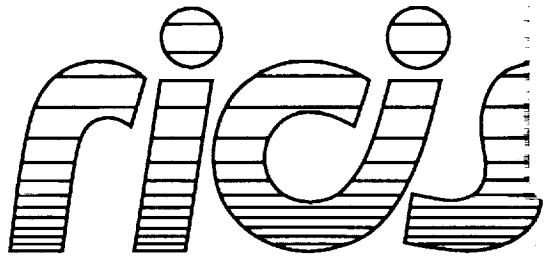
The NASA Visual Thesaurus: A Visual Interface to Visual Data

M.E. Rorvig
H. Turner
J. Moncada

Project Icon Image Scaling Laboratory
The University of Texas at Austin

Cooperative Agreement NCC 9-16
Research Activity No. IM.10

NASA Johnson Space Center
Center Operations Directorate
Photography and Television Technology Division



Research Institute for Computing and Information Systems
University of Houston-Clear Lake

TECHNICAL REPORT

The RICIS Concept

The University of Houston-Clear Lake established the Research Institute for Computing and Information Systems (RICIS) in 1986 to encourage the NASA Johnson Space Center (JSC) and local industry to actively support research in the computing and information sciences. As part of this endeavor, UHCL proposed a partnership with JSC to jointly define and manage an integrated program of research in advanced data processing technology needed for JSC's main missions, including administrative, engineering and science responsibilities. JSC agreed and entered into a continuing cooperative agreement with UHCL beginning in May 1986, to jointly plan and execute such research through RICIS. Additionally, under Cooperative Agreement NCC 9-16, computing and educational facilities are shared by the two institutions to conduct the research.

The UHCL/RICIS mission is to conduct, coordinate, and disseminate research and professional level education in computing and information systems to serve the needs of the government, industry, community and academia. RICIS combines resources of UHCL and its gateway affiliates to research and develop materials, prototypes and publications on topics of mutual interest to its sponsors and researchers. Within UHCL, the mission is being implemented through interdisciplinary involvement of faculty and students from each of the four schools: Business and Public Administration, Education, Human Sciences and Humanities, and Natural and Applied Sciences. RICIS also collaborates with industry in a companion program. This program is focused on serving the research and advanced development needs of industry.

Moreover, UHCL established relationships with other universities and research organizations, having common research interests, to provide additional sources of expertise to conduct needed research. For example, UHCL has entered into a special partnership with Texas A&M University to help oversee RICIS research and education programs, while other research organizations are involved via the "gateway" concept.

A major role of RICIS then is to find the best match of sponsors, researchers and research objectives to advance knowledge in the computing and information sciences. RICIS, working jointly with its sponsors, advises on research needs, recommends principals for conducting the research, provides technical and administrative support to coordinate the research and integrates technical results into the goals of UHCL, NASA/JSC and industry.

The NASA Visual Thesaurus: A Visual Interface to Visual Data

Preface

This research was conducted under auspices of the Research Institute for Computing and Information Systems by Dr. Mark E. Rorvig, Dr. Ronald Willis, H. Turner and J. Moncada of the Project Icon Scaling Laboratory, Graduate School of Library and Information Science at the University of Texas at Austin. Dr. Peter C. Bishop, Director of the Space Business Research Center, University of Houston-Clear Lake served as RICIS research coordinator.

Funding has been provided by the Center Operations Directorate, NASA/JSC through Cooperative Agreement NCC 9-16 between the NASA Johnson Space Center and the University of Houston-Clear Lake. The original NASA technical monitor for this research activity was Paul R. Penrod, Assistant to the Director, Center Directorate. Upon his retirement, David G. Billingsley, Chief, Photography and Television Technology Division, Center Operations Directorate, NASA/JSC became the technical monitor.

The views and conclusions contained in this report are those of the authors and should not be interpreted as representative of the official policies, either express or implied, of NASA or the United States Government.

THE NASA VISUAL THESAURUS: A VISUAL INTERFACE TO VISUAL DATA

M.E. RORVIG, H. TURNER, J. MONCADA
PROJECT ICON IMAGE SCALING LABORATORY
GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712-1276

DESCRIPTION OF THE PROJECT

Civilization's acknowledgement of the value of images to support myriad cognitive processes is recorded in the homely phrase "A picture is worth a thousand words." Increasingly, technical support and documentation processes rely on images to convey meaning on complex issues [1,2]. Commensurate with this trend is the rapid growth of institutional image collections (which both support and augment technical goals) since the mid-1950's. Unfortunately, success at creating collections of images has far exceeded the ability to access, describe, and manage them.

One of the more important image collections to be assembled during the last thirty years is the photographic archive of the National Aeronautics and Space Administration maintained through the Photography and Television Technology Division at the Johnson Space Center in Houston, Texas. This collection consists of over a million still photographs as well as several thousand films and includes imagery made available through NASA's Public Affairs Office and the National Aeronautics and Space Museum. Since October of 1987, the Project ICON Image Scaling Laboratory at The University of Texas at Austin has been engaged in the creation of a controlled vocabulary by which this collection could be indexed and searched. This effort has been conducted using methods of automatic thesaurus construction first explored by Salton and others at Cornell.

Early on, however, it became apparent that a traditional, printed thesaurus would be quite difficult to introduce into the existing processes of indexing and searching at the Johnson Space Center without severe increases in cost. Moreover, due to the high degree of inconsistency observed among indexers asssigning terms to images in any environment, it was recognized that many of the major benefits of term disambiguation dervied from thesauri would probably fail to be realized. This state of affairs led the ICON Laboratory to propose a new form of thesuarus for indexing and searching this visual collection. Called "The Visual Thesuarus", because of its equivalencing of images drawn from the collection with broad catagories of terms, this tool is thought to be applicable to many different collections of imagery far afield from the domain of aerospace ventures alone. This paper describes the background ideas for this project. A prototype will be made available in succeeding months from the authors.

The central technological problem addressed by this project is access to all visual content databases. Examples of such databases are: 1) Radiological and pathological collections in hospitals, medical schools and public health facilities; 2) Engineering drawings in goverment and private industry; 3) Architectural drawings and plans; 4) Geographic feature maps in geologic overhead (e.g. LANDSAT) imagery; 4) Photographic documentation collections in contruction and engineering; 5) General photography archives and collections. Specific access problems addressed are: 1) The accurate and rapid description of images entering collections; 2) Searching of collections of images.

The existing framework of image retrieval technology is adequately reviewed by S.K. Chang [3] and L.F. Lunin [4] with a reasonable description of a prototypical system available in Gale [5]. The market for image processing systems and components is a large one and is expected to grow from $415 million in 1986 to $1.6 billion by 1990 [**Frost and Sullivan, Inc.**]. However, one of the major impediments to the successful growth and product introduction processes in this industry remains the difficulty of describing and retrieving images in visual

collections. The summary paragraph by Lunin describing the present state of requirements in image information systems is of interest regarding this research and is quoted below **[4, p. 207-8]**.

> "The need has been proclaimed. The hardware and software are being developed. Standards are being discussed. The field is ready for unifying concepts. Yet one of the main areas of concern is how to represent structural information in a database mode... The designs and techniques for image databases still seem to have a long way to go before users can retrieve by 'language of the image' and browse among these databases easily and efficiently."

The description and searching of images are closely related processes. Without accurate description, there cannot be accurate retrieval. Therefore, an important goal of this project is the creation of a production system for the rapid and accurate description of images in the "language of the image" applicable to various image collection domains. Two problems must be addressed in this effort. First, image description is a time consuming process, often requiring high-level personnel for accurate results **[6]**. Image description practices are therfore often "sparse" in that descriptions are not sufficiently rich in linguistic content to support the variety of inquiries posed against them [7]. Second, individual differences in image perception give rise to extraordinary idiosyncrasy in the assignment of image terms. Markey, for example, notes that three indexers assigning terms to thirteen images exhibited only a 13% overlap in descriptive term assignment **[8]**. As noted earlier, conventional efforts to impose order on such term assignment variations usually include the construction of thesauri to guide indexers in the choice of terms. But such instruments are difficult to use, and, in the case of images, are unlikely in themselves to eliminate the ideosyncratic response so prevalent in human reactions to visual materials.

The second goal of this project concerns the problem of searching visual content collections by "language of the image." In addition to (if not caused by) the problem of inexact, ideosyncratic, and sparse vocabulary assignments to images, there exist significant problems in image retrieval. It is, in fact, usually the case that retrieval of images is a frustrating, time-consuming and highly inexact task **[7]**. Thus, this project seeks the creation of a visual interface for searching visual collections. Such an interface is proposed to facilitate searching by permitting search arguments to be made against image databases by selecting images from a display screen which correspond to linguistic descriptions. The images used to represent the linguistic terms, upon selection, may simply be converted internally by the interface system to their corresponding linguistic strings; and those strings in turn submitted as search arguments to conventional text-oriented retrieval systems in either partial match or Boolean logic implementations.

ILLUSTRATION A depicts a conventional process for the indexing and retrieval of images. **ILLUSTRATION B** on the same page depicts the proposed system of visual interfaces for visual content database description and searching. In general, the significant expectation of the visual interface, or thesaurus of searching exemplars, is a reduction of the ambiguity (i.e, guesswork) involved with term assignment in the indexing process and in the searching process. However, although there is abundant evidence that this intended benefit would be derived from this proposed form of interface **[6,9,10,11,12]**, the procedures to create the interface and the technology to operationalize the interface have only recently become capable of simultaneous interaction. The following paragraphs thus describe the intellectual procedures necessary to support the creation of the interface as well as the technology

necessary to operationalize the interface.

The fundamental intellectual problem posed by the creation of the visual interface is the accurate matching of representative images to term classes. There are two separate processes available to achieve this objective. The first of these, as summarized in [13], is purely statistical, has been well-demonstrated in laboratory environments for medical collections, and is chiefly applicable to collections of images with existing linguistic descriptive terms. The second process [14], involves both statistical analysis and direct, psychometric observations of human judgments of the similarity between images and linguistic expressions. This process is chiefly applicable to collections of images without significant existing textual descriptions, i.e., sparsely described collections. Neither of these techniques has ever been applied to the process of matching images to terms and must be considered tentative. Thus, a third goal of this project is the development of a set of standard procedures which may be applied to any domain in which a visual interface would be viewed as beneficial.

ILLUSTRATION C provides a description of the process adapted from [13] and proposed for use in the task of matching images to terms for collections of images already conjoined with textual descriptions. As noted in this illustration, the procedure involves the creation of a term association matrix to establish broad term hierarchies, and the subsequent inversion of the matrix to locate cluster centroids: those centroid image identification numbers referring to the images best representative of the corresponding class of terms. It should be noted that Illustration C rather simplifies these procedures and interested readers should consult the chapter references listed in [13] for further details.

ILLUSTRATION D describes a technique applicable to situations in which items in visual collections are described by a small number of broad classes. The goal of the procedure is to find the best match between "image exemplars" and each broad class, i.e., to find among those images assigned to a class, the one most indicative and evocative of the attributes of that class. The statistical methods of Illustration C and reference [13] are inapplicable in this case, because the probability density space of terms surrounding the images is simply too constrained. Thus, in this procedure, psychometric methods are proposed for use in completing the process of matching representative images to broad classes. This technique was applied first in [14] some years ago in the context of matching index terms to documents in document description and retrieval processes. Renewed interest in this work is noted in [15].

Although this latter procedure seems quite labor intensive, it must be noted that the actual judgments may be taken quite quickly. For example, for 250 broad classes, with an average of fifteen images to be selected for each class at approximately five seconds per judgment, the fifteen individuals chosen as selectors can provide this information in about one week of concentrated time or in about three weeks scheduled at convenient intervals.

The related tasks of this project concern the construction of the "interface engine" itself, as well as the communications interfaces necessary for a user transparently to inquire of text oriented systems in a visual mode. Most of the hardware and software required for these tasks has only recently become available in "off-the-shelf" versions, but it should be noted that all the technology necessary to implement the proposed visual interface is presently available. The development effort is currently in progress within the Apple Computer MacIntosh II environment, specifically employing HyperCard as the relational database tool. Much of the Capacities of the visual

thesuarus however may also be transportable to the monochrome environments of the MacIntosh SE and the MacIntosh + though with some degradation of speed. Although there is much debate possible over the various merits of HyperCard itself, as well as closely related software products, the development of the intellectual tools for constructing visual interfaces need not be restricted to HyperCard, or the MacIntosh II environment at all. This environment has been selected because (1) there is wide agreement that these types of products will be widely used throughout the next five years, (2) that HyperCard is a good exemplar of software products which will be developed in the future, and (3) that products developed in this environment can be distributed easily and cheaply through hundreds of commercial outlets as so called "HyperCard Stacks".

ILLUSTRATION E details the types of relations existing presently among vocabularies consisting of strings, and their presentation as images in a HyperCard environment. In this environment, system indexers and searchers may easily flip back and forth between images and texts, with the central goal to permit the use of the visual representation to reduce ambiguity in term assignment in both indexing and searching procedures. **ILLUSTRATION F** displays a prototype of the interface requirements which a system must have in order to search existing (in this example Boolean search oriented) textual databases of image description "in the language of the image" (see also **ILLUSTRATION B.**)

Two collections of visual content materials are presently in use for this development. The first of these is applicable to the general NASA photographic archives. This collection consists of about one million photographs created since the early 1960's and covering every aspect of NASA tasks since that period. Initially, a sample of approximately 10,000 images will be used for this project. These images have been selected on the basis of high use and strategic interest by NASA photographic collection managers. The descriptions for each photo range from about 25 to 100 words and will be treated by the methods described in **ILLUSTRATION C.** The second collection is from the Earth Observations Division of the NASA. This collection consists of about 50,000 photographs from the hand held photography program obtained on various space shuttle missions. Descriptions of these images are "sparse" and consist of assignments of single terms from a vocabulary of about 2000 terms. This collection will be treated by the method described in **ILLUSTRATION D.**

Several evaluation tests of the developed interfaces are proposed. However, the evaluation techniques are quite straightforward. Since both collections noted above and used in this research are working collections, that is, collections presently in use by researchers, improvements in indexing and searching efficiency may be measured in terms of minutes and hours. The experimental design for the evaluation tests are available in **ILLUSTRATION G**, with a quantitative expression of the underlying theory available in **[16].**

The resources presently in use for this research include the computing facilities provided by The University of Texas at Austin as well as the facilities of the Project ICON Image Scaling Laboratory. The Project ICON Laboratory has been supported by grants from the The Council on Library Resources, IBM Corporation, DataCopy Corporation, The University Research Institute of The University of Texas at Austin, and NASA. The Project ICON Laboratory is one of a number of facilities set up to conduct research on aspects of intellectual access to graphic records. Other similar laboratories exist at MIT, UC-Berkeley, and The University of Pittsburg.

THE NASA VISUAL THESAURUS: A VISUAL INTERFACE TO VISUAL DATA
M.E. RORVIG, H. TURNER, J. MONCADA
PROJECT ICON IMAGE SCALING LABORATORY
GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712-1276

(I) REFERENCES

[1] Lesk, M. E. Graphical information resources: Maps and beyond. Proceedings of the Eighth Annual International Conference on Research and Development in Information Retrieval. Association for Computing Machinery; New York; 1985:1-8.

[2] National Center for Supercomputer Applications. Scientific visualization program announcement. University of Illinois; Champaign, Illinois; 1987.

[3] Chang, S.K. Image database systems. In Handbook of Pattern Recognition and Image Processing (T.Y. Young and K.S. Fu, eds.). Academic Press; Orlando, Florida;1986:371-395.

[4] Lunin, L.F. Electronic image information. In Annual Review of Information Science and Technology Volume 22, 1987 (M. E. Williams, ed.). Elsevier Science; Amsterdam;1987:179-224.

[5] Gale, J.C. Use of optical disks for information storage and retrieval. Information Technology and Libraries. 3(4); 1984:379-382.

[6] Rorvig, M.E. The substitutibility of images for textual description of archival materials in an MS-DOS environment. In The Application of Micro-Computers in Information, Documentation, and Libraries (K. Lehman and H. Strohl-Goebel, eds.). North-Holland; Amsterdam;1987:407-415.

[7] Zappert, F.G. Personal communication of February 4, 1988 describing Dialog Information Servies, Inc. efforts to construct an indexing and searching facility for trademark access. Advanced Technology Group. Dialog Information Services, Inc.; Palo Alto, CA 94304.

[8] Markey, K. Interindexer consistency tests: A literature review and report of a test of consistency in indexing visual materials. Library and Information Science Research. 6;1984:155-177.

[9] Wyant, S.; Banks, W. P.; Berger, D. Verbal and pictorial similarity in recognition of pictures. Perception and Psychophysics. 12(2a);1972:151-153.

[10] Golden, R. N. Pictures at a Recognition: Multidimensionally Scaled Pictures and Forced-Choice Recognition Performance. Unpublished Ph.D. Dissertation. Department of Psychology, University of Washington; Seattle:1984.

[11] Walters, D.K.W. A computer vision model based on psychophysical experiments. In Pattern Recognition by Humans and Machines, Volume 2 (E. C. Schwab and H.C. Nusbaum, eds.). Academic Press; Orlando, Florida; 1986:88-120.

[12] Tamura, H.; Mori, S; Yamawaki, T. Textural features corresponding to visual perception. IEEE Transactions on Systems, Man, and Cybernetics. 8;1978:460-473.

[13] Salton, G.; McGill, M. J. Introduction to Modern Information Retrieval. McGraw-Hill; New York:1983.

[14] Katter, R. V. Study of Document Representations: Multidimensional Scaling of Indexing Terms. Technical Memorandum. Systems Development Corporation; Santa Monica, California: 1967.

[15] Rorvig, M. E. Psychophysics and information retrieval. In Annual Review of Information Science and Technology, Volume 23, 1988 (M. E. Williams, ed.). Elsevier Science; Amsterdam:[in progress].

[16] Harmon, E. G. The measurement of information. Information Processing and Management. 20;1984:193-198.

[17] Harter, S. P. A probabilistic approach to keyword indexing, Part 1: On the distribution of specialty words in a

THE NASA VISUAL THESAURUS: A VISUAL INTERFACE TO VISUAL DATA

M.E. RORVIG, H. TURNER, J. MONCADA
PROJECT ICON IMAGE SCALING LABORATORY
GRADUATE SCHOOL OF LIBRARY AND INFORMATION SCIENCE
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712-1276

technical literature. Journal of the American Society for Information Science. 26(4);1975:197-206.

ILLUSTRATION A:
CONVENTIONAL IMAGE DESCRIPTION AND RETRIEVAL

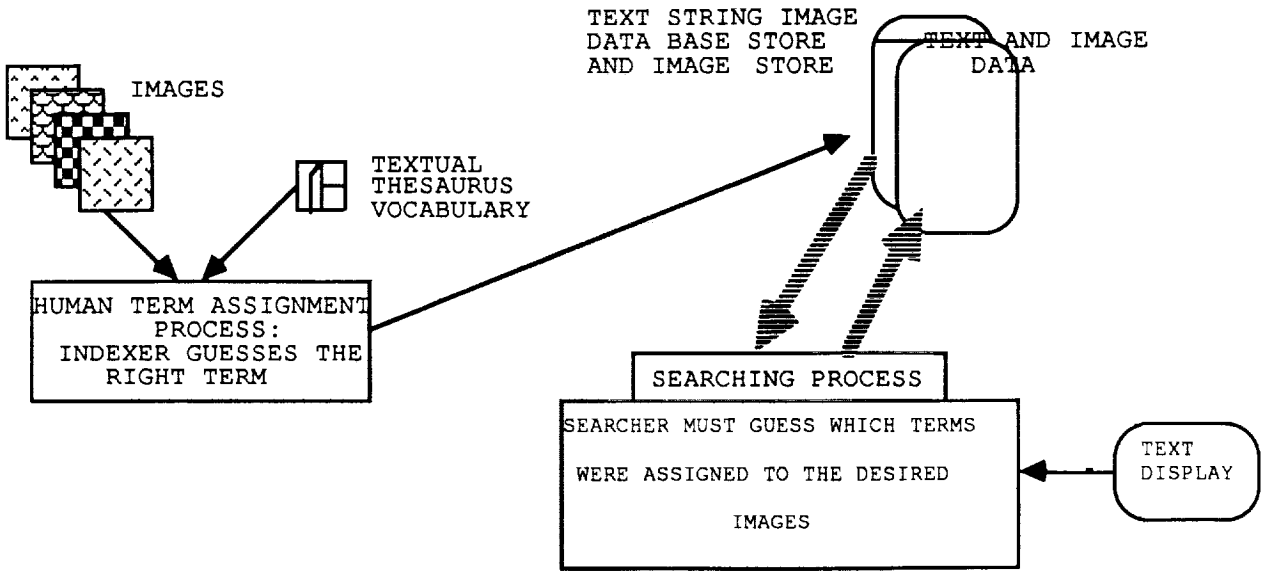


ILLUSTRATION B:

PROPOSED VISUAL THESAURUS INDEXING AND RETRIEVAL SYSTE

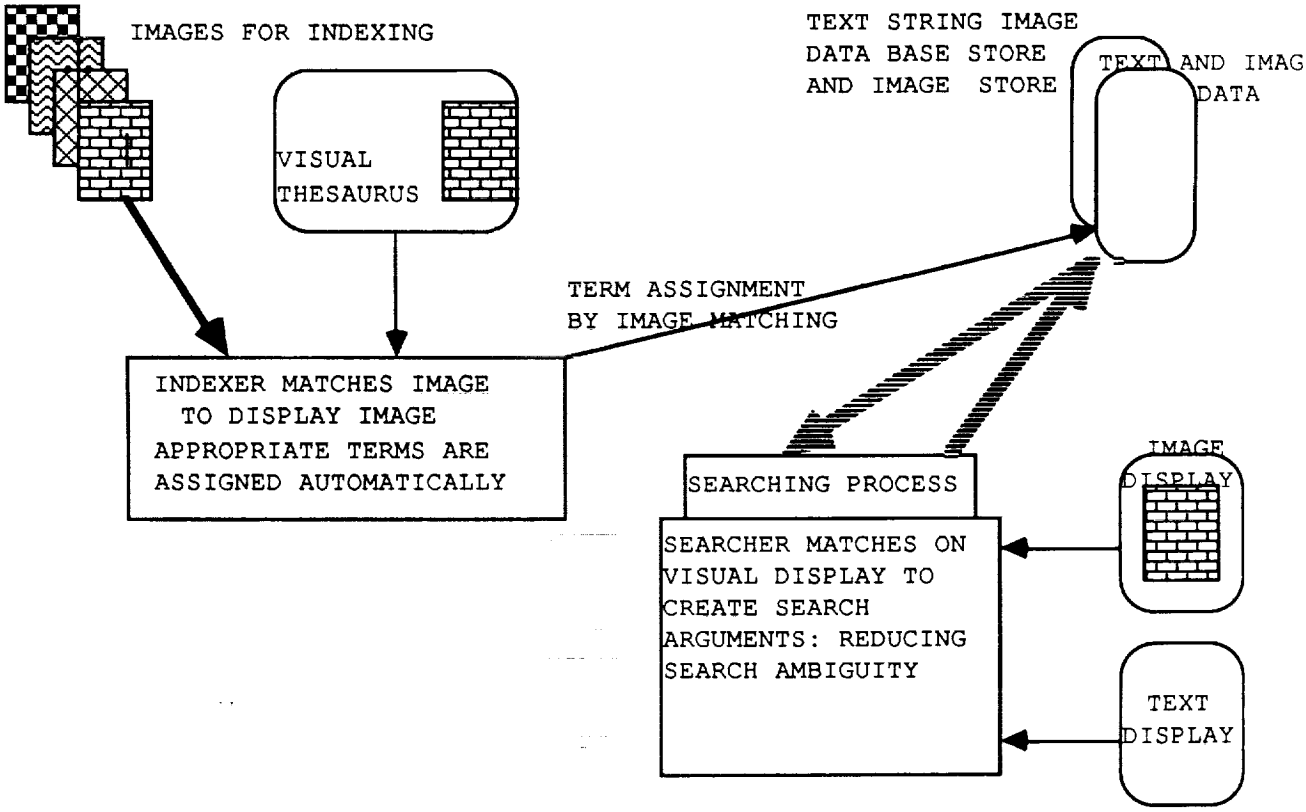


ILLUSTRATION E:
THESAURUS TERMS IN STRING AND VISUAL COMPONENTS

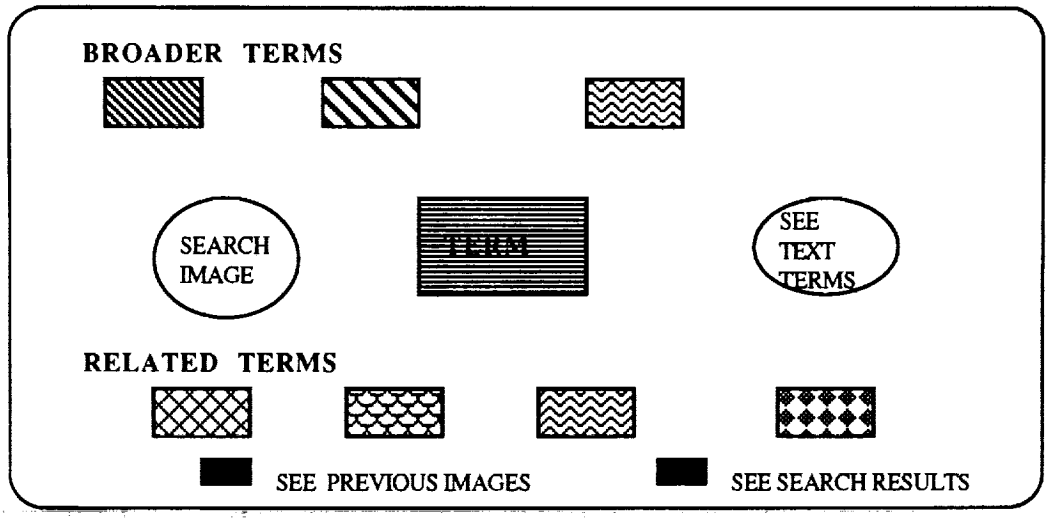
TEXTUAL VERSION EXAMPLE

GEOLOGICAL FAULTS

(Broader terms)
RIFTS
SPLITS
THRUST FAULTS

(Related terms)
CRUSTAL FRACTURES
FISSURES
FOLDS
FORMATIONS

HYPERCARD EXAMPLE

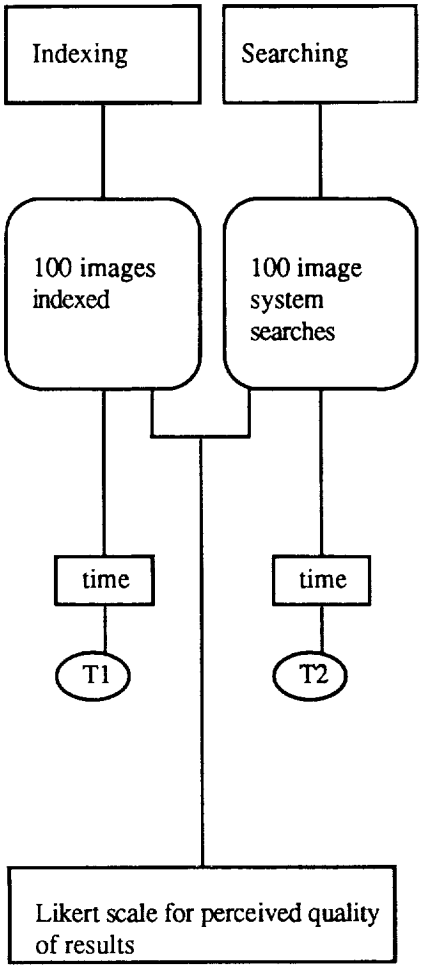


UNLIKE THE TEXTUAL EXAMPLE ABOVE, SEARCHERS MAY VIEW A VISUAL REPRESENTATION OF THE EXACT TERM, MAKE A SEARCH, SEE THE RESULTS OF A SEARCH, OR LOOK AT IMAGES PREVIOUSLY SELECTED.! MOREOVER, BROADER TERMS AND NARROWER TERMS MAY BE EXPANDED BY PRESSING THE MOUSE BUTTON TO SEE OTHER TERM RELATION SETS IN VISUAL FORM.

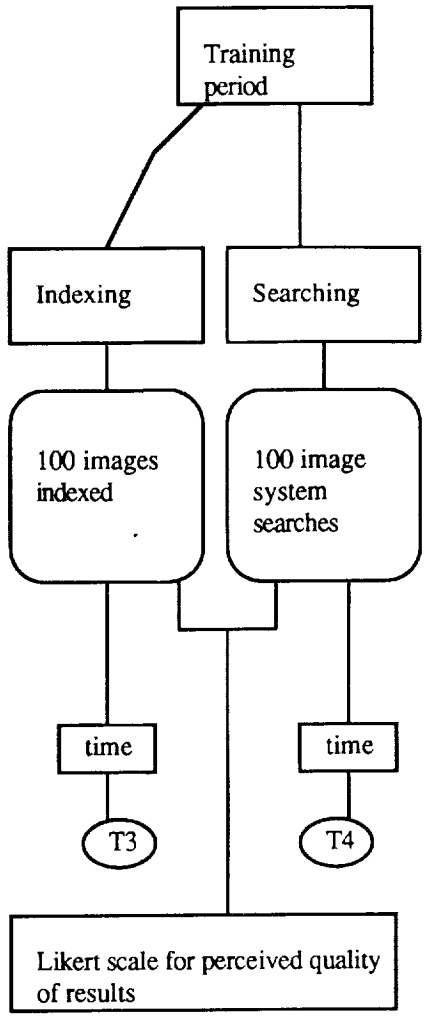
ILLUSTRATION G:

SYSTEM EVALUATION

CURRENT SYSTEMS
NASA-JSC / GALVESTON



PROPOSED VISUAL SYSTEM



Because both of the test collections used in this proposed research are working collections, it will be possible to test the system by introducing it into the working environments for indexing images and searching images. Assuming two time period samples free from biasing factors (such as special work requirements, etc.), the methods presently in use may be timed in minutes and hours for 100 tasks, the new visual tool introduced, and 100 tasks timed again. A successful outcome is thus (T1 + T2) < (T3+T4) with similar Likert scale ratings for perceived quality of results.